Project 2

STATS 220 Fall 2025

Your Name

## Overview

I examined a dataset derived from my personal food delivery orders over a specific time frame for this project. The dataset includes details on the various ordering platforms, restaurant types, total order amounts, delivery schedules, and customer reviews. My objective was to investigate important elements like ratings, delivery times, and the relationship between platform selections and client satisfaction.

My personal food delivery orders were tracked and stored in an easily accessible format in order to gather the data. To ensure a comprehensive picture of my delivery experiences, the collection process concentrated on the platforms used, restaurant types, order amounts, delivery times, and ratings. The information was gathered with the least amount of bias and represents actual experiences.

When creating the form, I concentrated on simplicity and clarity ensure accurate and speedy data entry. It was simple to choose the delivery platform, restaurant type, and delivery rating thanks to the form’s options. This made sure that the information recorded accurately represented my actual food delivery experiences.

I discovered patterns about which platforms provided the quickest delivery times and the highest customer satisfaction scores by examining the data. The analysis was a helpful guide for future food delivery decisions and gave me a better understanding of how well various delivery platforms carried out.

<https://docs.google.com/forms/d/e/1FAIpQLSfIOIYx1OmHC5Eg586IhuwV-F_kOfQBvHXUWiafQuyO7w2Olg/viewform?usp=header>

##Dynamic report

<https://github.com/wany840/stats220/blob/main/project-2.docx>

Because my computer for storing documents was stolen (you can check the video link I sent),

<https://drive.google.com/file/d/1PyhGnbQbTj4S_5Hew74OZa_sewa0S84f/view?usp=sharing> I could only write these documents on my phone. I set up everything including checking the operation on R language and found that it could not connect to the domain, so I uploaded the resulting word text as a compensation.

##Creativity

I made some additional innovations, i.e., find whose delivery time is shortest, or whose food experience is better, when the dishes I want to eat are set and the time is set (which is also a set requirement for most people). But I cannot make a detailed analysis due to a lack of data.

##Learning reflection: technical competence and breakthrough of thinking in dynamic reporting system I have derived twin inspiration in learning of dynamic reporting system from this study. On a technical level, I have learned the main process chain of creating dynamic reports: collect information through Google Forms → automatically maintain it in Google Sheets → deploy as CSV → analyze and process with R language → generate dynamic reports. Particularly, I learned how to utilize read\_csv() to read data on the web in real time, and use paste() function to insert analysis results dynamically into report text.

At the thought level, this study completely changed my understanding of the nature of reports. I understood that a good reporting system ought to have three features: first, “breathing” - it is able to grow organically with the injection of new data; second, “dialogue” - establish two-way conversation with readers using interactive features; and finally, “narrative power” - leave the story of the data open and liable to change. This change of mind made me start looking at data analysis systems from the life form angle.

What excites me most is that these technologies can be extended to more innovative areas:

Intelligent early warning system: trigger automatic alerts by setting thresholds

Personalized reports: dynamically adjust content presentation based on user characteristics

Collaborative analysis platform: multiple people edit and verify data in real time

Augmented reality reports: visualize dynamic data in a physical environment

I especially want to explore how to integrate natural language processing technology into dynamic reports to achieve:

Automatically generated data interpretation copy

Smart question and answer interactive feature

Multi-language real-time translation

These advanced applications will greatly improve the intelligence level of data analysis and user experience.

##Appendix

---  
  
title: "exploration.R"  
  
date: "2025-03-30"  
---  
  
```{r setup, include=FALSE}  
  
library(tidyverse)  
```  
  
## R Markdown  
  
This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.  
  
When you click the \*\*Knit\*\* button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:  
  
```{r }  
# Load necessary libraries  
library(dplyr)  
library(ggplot2)  
library(corrplot)  
  
# Load the data  
logged\_data <- read.csv("foodsurvey.csv")  
  
# Rename columns for easier understanding  
latest\_data <- logged\_data %>%  
 rename(  
 Platform = `Which.food.delivery.platform.was.used.for.this.order.`,  
 RestaurantType = `What.type.of.restaurant.was.this.order.from.`,  
 TotalAmount = `What.was.the.total.amount.of.the.order...Currency..NZD.`,  
 OrderTime = `Order.Time..Multiple.Choice.`,  
 DeliveryTime = `How.many.minutes.did.it.take.from.ordering.to.delivery.`,  
 DeliveryAttitude = `How.was.the.delivery.person.s.attitude.`,  
 Rating = `X...How.would.you.rate.this.food.delivery.experience...1...Very.poor..10...Excellent...`  
 )  
  
# ---- Descriptive Statistics ----  
  
# 1. Calculate the mean, median, and standard deviation of TotalAmount and DeliveryTime  
mean\_total\_amount <- mean(latest\_data$TotalAmount, na.rm = TRUE)  
median\_total\_amount <- median(latest\_data$TotalAmount, na.rm = TRUE)  
sd\_total\_amount <- sd(latest\_data$TotalAmount, na.rm = TRUE)  
  
mean\_delivery\_time <- mean(latest\_data$DeliveryTime, na.rm = TRUE)  
median\_delivery\_time <- median(latest\_data$DeliveryTime, na.rm = TRUE)  
sd\_delivery\_time <- sd(latest\_data$DeliveryTime, na.rm = TRUE)  
  
# Print results  
cat("Mean Total Amount:", mean\_total\_amount, "\n")  
cat("Median Total Amount:", median\_total\_amount, "\n")  
cat("Standard Deviation of Total Amount:", sd\_total\_amount, "\n")  
  
cat("Mean Delivery Time:", mean\_delivery\_time, "\n")  
cat("Median Delivery Time:", median\_delivery\_time, "\n")  
cat("Standard Deviation of Delivery Time:", sd\_delivery\_time, "\n")  
  
```  
  
## Including Plots  
  
You can also embed plots, for example:  
  
```{r }  
# ---- Explore the data: Summary statistics ----  
  
# 1. Summary statistics for TotalAmount  
min\_total\_amount <- min(latest\_data$TotalAmount, na.rm = TRUE)  
max\_total\_amount <- max(latest\_data$TotalAmount, na.rm = TRUE)  
mean\_total\_amount <- mean(latest\_data$TotalAmount, na.rm = TRUE)  
median\_total\_amount <- median(latest\_data$TotalAmount, na.rm = TRUE)  
  
# 2. Summary statistics for DeliveryTime  
min\_delivery\_time <- min(latest\_data$DeliveryTime, na.rm = TRUE)  
max\_delivery\_time <- max(latest\_data$DeliveryTime, na.rm = TRUE)  
mean\_delivery\_time <- mean(latest\_data$DeliveryTime, na.rm = TRUE)  
median\_delivery\_time <- median(latest\_data$DeliveryTime, na.rm = TRUE)  
  
  
  
# ---- Visualizations ----  
  
# 1. Bar chart for the distribution of Food Delivery Platforms  
ggplot(latest\_data, aes(x = Platform)) +  
 geom\_bar(fill = "#B2C248") +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1)) +  
 labs(title = "Distribution of Food Delivery Platforms", x = "Platform", y = "Count")  
  
# 2. Bar chart for the distribution of Restaurant Types  
ggplot(latest\_data, aes(x = RestaurantType)) +  
 geom\_bar(fill = "#6a0dad") +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1)) +  
 labs(title = "Distribution of Restaurant Types", x = "Restaurant Type", y = "Count")  
  
# 3. Boxplot showing relationship between Delivery Time and Platform  
ggplot(latest\_data, aes(x = Platform, y = DeliveryTime)) +  
 geom\_boxplot(fill = "#6a0dad") +  
 labs(title = "Delivery Time by Platform", x = "Platform", y = "Delivery Time (Minutes)") +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1))  
  
# 4. Bar chart for average Rating by Platform  
ggplot(latest\_data, aes(x = Platform, y = Rating)) +  
 stat\_summary(fun = mean, geom = "bar", fill = "#4CAF50") +  
 labs(title = "Average Rating by Platform", x = "Platform", y = "Average Rating") +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1))  
  
# 5. Scatter plot to check if higher price correlates with higher ratings  
ggplot(latest\_data, aes(x = TotalAmount, y = Rating)) +  
 geom\_point(alpha = 0.5, color = "#FF5733") +  
 labs(title = "Price vs Rating", x = "Total Amount (NZD)", y = "Rating")  
  
# ---- Comments Section ----  
  
# 1. The platform with the fastest delivery time is identified.  
# 2. The platform with the highest average rating is identified.  
# 3. The highest rating based on platform, restaurant type, and order time is found.  
# 4. The fastest delivery times are identified for specific platform and restaurant type combinations.  
# 5. The effect of high price (>= 50 NZD) on ratings is analyzed.  
  
# Statistical insights:  
# - Higher priced food (>= 50 NZD) generally shows a higher average rating compared to lower priced food.  
# - The fastest platform and restaurant type combination can help optimize the delivery process.  
  
```"  
output: html\_notebook  
---  
  
This is an [R Markdown](http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.   
  
Try executing this chunk by clicking the \*Run\* button within the chunk or by placing your cursor inside it and pressing \*Ctrl+Shift+Enter\*.   
  
```{r}  
plot(cars)  
```  
  
Add a new chunk by clicking the \*Insert Chunk\* button on the toolbar or by pressing \*Ctrl+Alt+I\*.  
  
When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the \*Preview\* button or press \*Ctrl+Shift+K\* to preview the HTML file).  
  
The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike \*Knit\*, \*Preview\* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

---  
title: My dynamic report  
output: html\_fragment  
---  
  
```{r setup, include=FALSE}  
knitr::opts\_chunk$set(echo=FALSE, message=FALSE, warning=FALSE, error=FALSE)  
  
```  
  
```{css}  
  
```  
  
  
##Overview of the Data  
In this report, we are analyzing a dataset of food delivery orders. The dataset contains information on different platforms used for ordering, the types of restaurants, total order amounts, and customer ratings.  
  
We aim to explore key factors like delivery times, ratings, and the correlation between platform choices and customer satisfaction.  
  
##Delivery Time Analysis  
We begin by analyzing the average delivery time across different food delivery platforms.  
  
```{r}  
library(dplyr)  
library(ggplot2)  
  
library(dplyr)  
library(ggplot2)  
  
logged\_data <- read.csv("C:/Users/wany840/Downloads/Project 2/foodsurvey.csv")  
  
# Rename columns for easier understanding  
latest\_data <- logged\_data %>%  
 rename(  
 Platform = `Which.food.delivery.platform.was.used.for.this.order.`,  
 RestaurantType = `What.type.of.restaurant.was.this.order.from.`,  
 TotalAmount = `What.was.the.total.amount.of.the.order...Currency..NZD.`,  
 OrderTime = `Order.Time..Multiple.Choice.`,  
 DeliveryTime = `How.many.minutes.did.it.take.from.ordering.to.delivery.`,  
 DeliveryAttitude = `How.was.the.delivery.person.s.attitude.`,  
 Rating = `X...How.would.you.rate.this.food.delivery.experience...1...Very.poor..10...Excellent...`  
 )  
  
# ---- Descriptive Statistics ----  
  
# Calculate the mean, median, and standard deviation of TotalAmount and DeliveryTime  
mean\_total\_amount <- mean(latest\_data$TotalAmount, na.rm = TRUE)  
median\_total\_amount <- median(latest\_data$TotalAmount, na.rm = TRUE)  
sd\_total\_amount <- sd(latest\_data$TotalAmount, na.rm = TRUE)  
  
mean\_delivery\_time <- mean(latest\_data$DeliveryTime, na.rm = TRUE)  
median\_delivery\_time <- median(latest\_data$DeliveryTime, na.rm = TRUE)  
sd\_delivery\_time <- sd(latest\_data$DeliveryTime, na.rm = TRUE)  
  
# Average delivery time by platform  
ggplot(latest\_data, aes(x = Platform, y = DeliveryTime)) +  
 stat\_summary(fun = mean, geom = "bar", fill = "#FF5733") +  
 labs(title = "Average Delivery Time by Platform", x = "Platform", y = "Average Delivery Time (Minutes)") +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1))  
  
```  
  
  
Customer Rating Analysis  
Next, we analyze the average ratings across different platforms.  
  
```{r}  
# Average rating by platform  
ggplot(latest\_data, aes(x = Platform, y = Rating)) +  
 stat\_summary(fun = mean, geom = "bar", fill = "#4CAF50") +  
 labs(title = "Average Rating by Platform", x = "Platform", y = "Average Rating") +  
 theme(axis.text.x = element\_text(angle = 45, hjust = 1))  
  
```  
  
  
## Dynamic Commentary Based on Data  
  
```{r}  
library(dplyr)  
library(ggplot2)  
  
latest\_data <- logged\_data %>%  
 rename(  
 Platform = `Which.food.delivery.platform.was.used.for.this.order.`,  
 RestaurantType = `What.type.of.restaurant.was.this.order.from.`,  
 OrderTime = `Order.Time..Multiple.Choice.`,  
 DeliveryTime = `How.many.minutes.did.it.take.from.ordering.to.delivery.`,  
 Rating = `X...How.would.you.rate.this.food.delivery.experience...1...Very.poor..10...Excellent...`  
 ) %>%  
 mutate(  
 # simplify the name for example l+d= lunch + doordash  
 Time\_Type = paste0(  
 substr(OrderTime, 1, 1), # take the first letter（L/D）  
 "+",  
 substr(RestaurantType, 1, 1)   
 )  
 )  
  
# count average  
summary\_data <- latest\_data %>%  
 group\_by(Time\_Type, Platform) %>%  
 summarise(  
 avg\_delivery\_time = mean(DeliveryTime, na.rm = TRUE),  
 avg\_rating = mean(Rating, na.rm = TRUE),  
 .groups = "drop"  
 )  
  
  
small\_font\_theme <- theme(  
 axis.text = element\_text(size = 8),  
 axis.title = element\_text(size = 9),  
 plot.title = element\_text(size = 10),  
 legend.text = element\_text(size = 8),  
 legend.title = element\_text(size = 9)  
)  
  
# delivery time plot  
ggplot(summary\_data, aes(x = Time\_Type, y = avg\_delivery\_time, fill = Platform)) +  
 geom\_col(position = position\_dodge(0.8), width = 0.7) +  
 labs(title = "Avg Delivery Time by Platform",   
 x = "Meal Time + Cuisine Type (L/D+W/A/L)",   
 y = "Delivery Time (min)") +  
 small\_font\_theme  
  
# rating table  
ggplot(summary\_data, aes(x = Time\_Type, y = avg\_rating, fill = Platform)) +  
 geom\_col(position = position\_dodge(0.8), width = 0.7) +  
 labs(title = "Avg Rating by Platform",   
 x = "Meal Time + Cuisine Type (L/D+W/A/L)",   
 y = "Rating (1-10)") +  
 small\_font\_theme  
```  
  
  
  
  
##Conclusion  
In conclusion, this report provides a thorough analysis of the food delivery dataset. We analyzed delivery times and customer ratings across platforms. Based on the data, we identified the platform with the fastest delivery time and the highest customer satisfaction.